

ON THE EXISTENCE OF NON-CENTRAL WISHART DISTRIBUTIONS

EBERHARD MAYERHOFER

ABSTRACT. This paper deals with the existence issue of non-central Wishart distributions which is a research topic initiated by Wishart (1928), and with important contributions by e.g., Lévy (1937), Gindikin (1975), Shanbhag (1988), Peddada & Richards (1991). We present a new method involving the theory of affine Markov processes, which reveals joint necessary conditions on shape and non-centrality parameter. While Eaton's conjecture concerning the necessary range of the shape parameter is confirmed, we also observe that it is not sufficient anymore that it only belongs to the Gindikin ensemble, as is in the central case.

1. INTRODUCTION

The general non-central Wishart distribution $\Gamma(p, \omega; \sigma)$ on the cone S_d^+ of symmetric positive semi-definite $d \times d$ matrices is defined (whenever it exists) by its Laplace transform

$$\mathcal{L}(\Gamma(p, \omega; \sigma))(u) = (\det(I + \sigma u))^{-p} e^{-\text{tr}(u(I + \sigma u)^{-1}\omega)}, \quad u \in S_d^+, \quad (1.1)$$

where $p \geq 0$ denotes its shape parameter, $\sigma \in S_d^+$ is the scale parameter and the parameter of non-centrality equals $\omega \in S_d^+$. In the case that $\omega = 0$, $\Gamma(p; \sigma) := \Gamma(p, 0; \sigma)$ is called the central Wishart distribution, which had been introduced in 1928 by Wishart [11]. In 1937, Lévy [8] showed that $\Gamma(p; \sigma)$ on S_2^+ is not infinitely divisible for invertible σ , which means that for some sequence of shape parameters $p_k \downarrow 0$, $\Gamma(p_k; \sigma)$ cannot exist. Gindikin [5], Shanbhag [10] and Peddada & Richards [9]¹ subsequently showed that for non-degenerate σ ,

$$(\det(I + \sigma u))^{-p} \quad (1.2)$$

can only be the Laplace transform of a non-trivial probability measure for shape parameters p belonging to the Gindikin ensemble

$$\Lambda_d = \left\{ \frac{j}{2}, \quad j = 1, 2, \dots, d-2 \right\} \cup \left[\frac{d-1}{2}, \infty \right).$$

Aim of this work is to investigate this fundamental existence issue in the non-central case. We shall show:

Theorem 1.1. *Let $d \in \mathbb{N}$, $p > 0$, $\omega, \sigma \in S_d^+$. The following hold:*

Key words and phrases. non-central Wishart distribution, Gindikin ensemble, Wishart processes, Affine processes.

The author currently holds a Marie-Curie fellowship at Deutsche Bundesbank. The research leading to these results has received funding from the European Community Programme FP7-PEOPLE-ITN-2008 under grant agreement number PITN-GA-2009-237984 (RISK) and from the WWTF (Vienna Science and Technology Fund) The funding is gratefully acknowledged.

¹Contrary to [9] we exclude the point mass at zero, i.e. the Gindikin ensemble does not contain 0. Also, our notation deviates slightly from theirs, see Section A.

- (i) Suppose σ is invertible. If the right side of (1.1) is the Laplace transform of a non-trivial probability measure ² $\Gamma(p, \omega; \sigma)$ on S_d^+ , then $p \in \Lambda_d$ and $\text{rank}(\omega) \leq 2p + 1$.
- (ii) Conversely, suppose any of the following conditions hold:
 - (a) $p \geq \frac{d-1}{2}$,
 - (b) $p < \frac{d-1}{2}$ and $\text{rank}(\omega) \leq 2p$.
 Then the right side of (1.1) is the Laplace transform of a non-trivial probability measure $\Gamma(p, \omega; \sigma)$.

It should be noted that Theorem is not a full characterization of the existence of non-central Wishart distributions, because it leaves open the question, whether distributions $\Gamma(p, \omega; \sigma)$ exist with $p \in \{1/2, \dots, \frac{d-2}{2}\}$ and $\text{rank}(\omega) = 2p + 1$. This is only an interesting question for $d \geq 3$, and $\text{rank}(\omega) > 1$ as the following two corollaries demonstrate. These are immediate conclusions of Theorem 1.1.

Corollary 1.2. *Suppose that σ is invertible and $\text{rank}(\omega) \leq 1$. The following are equivalent:*

- (i) *The right side of (1.1) is the Laplace transform of a non-trivial probability measure $\Gamma(p, \omega; \sigma)$ on S_d^+ .*
- (ii) *$p \in \Lambda_d$.*

Another trivial consequence holds in low dimensions. Note that $\Lambda_1 = [0, \infty)$ and $\Lambda_2 = [\frac{1}{2}, \infty)$:

Corollary 1.3. *Let $d \leq 2$, and suppose σ to be invertible. The following are equivalent:*

- (i) *The right side of (1.1) is the Laplace transform of a non-trivial probability measure $\Gamma(p, \omega; \sigma)$ on S_d^+ .*
- (ii) *$p \in \Lambda_d$.*

We slightly adapt the notation of the recent article by Letac and Massam [7], and we are recollecting a number of fundamental statements thereof below (see section 2.1) especially what concerns basic properties of non-central Wishart distributions. Concerning their main statement as well as that of [9], the following important remark is due:

Remark 1.4. (i) [7, Proposition 2.3] claims that Λ_d fully characterizes the existence of non-central Wishart distributions. But this (paradoxically) allows the construction of Markovian Feller semigroups on S_d^+ which are non-positive, a mere impossibility. That's how we obtain the additional necessary conditions on the rank of ω in dependence of p , which suggests that the characterization of [7] is wrong. On the other hand, it is obvious that the existence proof of [7, Proposition 2.3, see also Proposition 2.1 and the subsequent paragraph] is incomplete, as for $p < \frac{d-1}{2}$ and $\text{rank}(\omega) > 2p$, the existence of non-central Wishart distributions $\Gamma(p, \omega; \sigma)$ is not shown there.

(ii) [9, Theorem 1] prove the necessity of $p \in \Lambda_d$ (which had been a conjecture by M.L. Eaton) under the premise that $\text{rank}(\omega) = 1$. However, the method of [9], which involves the theory of zonal polynomials, relies on the non-negativity of the so-called generalized binomial coefficients which may be "difficult to prove" in the case that $\text{rank}(\omega) > 1$, see their concluding remark in [9, Section 4]. In contrast, the present paper shows with a much simpler argument that $p \in \Lambda_d$, for *all* non-central Wishart distributions with nondegenerate scale parameter (see the first part of the proof of Theorem 1.1 (i)).

²It is easy to see that if $\sigma \neq 0$, the triviality of $\Gamma(p, \omega; \sigma)$ is equivalent to $p = 0$, and $\omega = 0$ (in which case we have the point mass at 0).

Our method for approaching the existence issue is new. We shall see that $p \in \Lambda_d$ can be proved by utilizing the situation in the central case (restated as Theorem 2.7) and Lévy's continuity theorem as well as a number of elementary facts such as (1) the behaviour of $\Gamma(p, \omega; \sigma)$ under the action of the linear automorphism group of S_d^+ and (2) the characterizing property of the natural exponential family associated with $\Gamma(p, \omega; \sigma)$, see Proposition 3.1. Concerning the rank condition, we proceed with an indirect argument: Assuming by contradiction that $p < \frac{d-1}{2}$ and $\text{rank}(\omega) > 2p+1$ allows the existence of affine Feller diffusions X supported on S_d^+ (see [2]) whose transition laws are non-centrally Wishart distributed. The contradiction $p \geq \frac{d-1}{2}$ is derived by observing that X violates a (geometric relevant) drift condition, as established in [2]. The latter is a consequence of the fact that the infinitesimal generator of a Markovian Feller semigroup satisfies the strong maximum principle, see [2, section 4.4].

1.1. Program of the paper. In section 2 we deliver notation and recall known facts about the existence of non-central Wishart laws (subsection 2.1) and Wishart processes on S_d^+ (subsection 2.2). The latter section uses a convenient notation for the Laplace transform, such that the distribution of Wishart processes can be easily read off from the characteristic exponents of the (affine) process, and which could be easily turned into an existence proof alternative to the one of [2]. The presentation of section 2.2 is instructive, and is of relevance for the proof of Theorem 1.1. Also, for the sake of completeness, we restate the characterization of the central Wishart distributions in terms of the Gindikin ensemble in subsection 2.3 and state trivial conclusions when σ is degenerate (characterization of existence and infinite divisibility). Section 3 presents a proof of Theorem 1.1. In Appendix A the relation of our definition of non-central Wishart distributions to others in the literature is given.

2. NOTATION AND PRELIMINARY RESULTS

Notation 2.1. Throughout the present article the following notation is relevant:

- \mathbb{R}_+ is the non-negative real line, and \mathbb{R}_{++} is its interior,
- M_d denotes the set of real $d \times d$ matrices, and S_d all symmetric ones therein.
- I is the unit $d \times d$ matrix.
- S_d^+ is the cone of symmetric positive semi-definite matrices, and S_d^{++} denotes its interior, the symmetric positive definite matrices. We denote its boundary $S_d^+ \setminus S_d^{++}$ by ∂S_d^+ .
- $\text{tr}(A)$ is the trace of a matrix $A \in M_d$, which introduces a scalar product on S_d via $\langle x, y \rangle := \text{tr}(xy)$ for $x, y \in S_d$.
- For $k = 1, 2, \dots, d-1$, $D_k \subseteq \partial S_d^+$ determines the (non-convex) cones of $d \times d$ matrices of rank less or equals k . We note that the dual cone of D_k equals $D_k^* = S_d^+$, hence the Laplace transform of a finite measure $\mu(d\xi)$ supported on D_k is defined by

$$\mathcal{L}(\mu)(u) := \int_{D_k} e^{-\langle u, \xi \rangle} \mu(d\xi), \quad u \in S_d^+.$$

2.1. Facts on non-central Wishart laws. First we recall the existence and basic properties of non-central Wishart distributions:

Lemma 2.2. *Let $p \in \Lambda_d$, $\sigma \in S_d^+$ and $\omega \in S_d^+$. We have:*

- Suppose $w = mm^\top$ for $m \in \mathbb{R}^d$ and set $\Sigma := \sigma/2$. If $Y \sim \mathcal{N}(m, \Sigma)$, then $X := YY^\top \sim \Gamma(1/2, w; \sigma)$ is supported on D_1 .*

- (ii) If $p < \frac{d-1}{2}$ and $\text{rank}(\omega) \leq 2p$, then the right side of (1.1) is the Laplace transform of a probability measure supported on D_{2p} .
- (iii) If $p \geq \frac{d-1}{2}$, then the right side of (1.1) is the Laplace transform of a probability measure $\Gamma(p, \omega; \sigma)$ on S_d^+ .
- (iv) In particular, if $p > \frac{d-1}{2}$ and if σ is invertible, then the density of $\Gamma(p, \omega; \sigma)$ exists³ and we denote it by $F(p, \omega, \sigma, \xi)$.

Proof. Proof of (i) is provided in [7, Proposition 3.2]. (ii): This follows from (i), by taking $2p$ independent normal random variables on \mathbb{R}^d with distribution $\mathcal{N}(m_i, \Sigma)$ where $\Sigma = \sigma/2$ and $\omega := m_1 m_1^\top + \dots + m_{2p} m_{2p}^\top$.

Note that if $\sigma \in S_d^{++}$, our definition of non-central Wishart distribution is related to the one of [7] in that $\Gamma(p, \omega; \sigma) = \gamma(p, \sigma^{-1} \omega \sigma^{-1}; \sigma)$, the latter being called "general non-central Wishart distribution" in [7] (see Appendix A for more detailed information). Hence statement (iv) is a consequence of [7, p. 1400].

Now for each $\varepsilon > 0$ we regularize σ and a by setting

$$\sigma_\varepsilon := \sigma + \varepsilon I, \quad a_\varepsilon := (\sigma + \varepsilon I)^{-1} \omega (\sigma + \varepsilon I)^{-1}.$$

Then for each $\varepsilon > 0$, we pick X_ε , an S_d^+ valued random variable according to [7, Proposition 2.3] such that

$$X_\varepsilon \sim \Gamma(p, \omega; \sigma_\varepsilon) (= \gamma(p, a_\varepsilon; \sigma_\varepsilon)).$$

Letting $\varepsilon \rightarrow 0$ and using Lévy's continuity theorem, we infer that X_ε converges in distribution to some random variable $X \sim \Gamma(p, \omega; \sigma)$. This settles part (iii) and (iv). \square

2.2. On the Fourier-Laplace transform of Wishart processes. A stochastically continuous Markov process $(X, \mathbb{P}_x)_{x \in S_d^+}$ on S_d^+ is called affine, if its Laplace transform is exponentially affine in the state-variable (see [2]). That is, for all $(t, x, u) \in \mathbb{R}_+ \times (S_d^+)^2$

$$\mathbb{E}[e^{-\langle u, X_t \rangle} \mid X_0 = x] = e^{-\phi(t, u) - \langle \psi(t, u), x \rangle}, \quad u \in S_d^+ \quad (2.1)$$

holds, where the so-called characteristic exponents ϕ and ψ satisfy a system of generalized Riccati equations,

$$\dot{\phi}(t, u) = F(\psi(t, u)), \quad \phi(0, u) = 0, \quad (2.2)$$

$$\dot{\psi}(t, u) = R(\psi(t, u)), \quad \psi(0, u) = u, \quad (2.3)$$

and F, R are of a specific Lévy-Khintchine form, which is particularly simple in the case of Wishart processes (which are pure diffusions; for the original definition in terms of stochastic differential equations and particular solutions of these SDEs, see [1]):

Definition 2.3. An affine process $X = (X, \mathbb{P}_x)_{x \in S_d^+}$ is a Wishart process on S_d^+ with parameter $(p \geq 0, \alpha \in S_d^+, \beta \in M_d)$, if its characteristic exponents (ϕ, ψ) satisfy the following Riccati equations:

$$\dot{\phi}(t, u) = 2p \langle \alpha, \psi(t, u) \rangle, \quad \phi(0, u) = 0 \quad (2.4)$$

$$\dot{\psi}(t, u) = -2\psi(t, u) \alpha \psi(t, u) + \psi(t, u) \beta + \beta^\top \psi(t, u), \quad \psi(0, u) = u. \quad (2.5)$$

Using notation and language from [2, Definition 2.3 and the discussion in section 2.1], the parameters satisfy the following

³For a detailed exposition of the densities, which involves the zonal polynomials, we refer to [7, eq. p. 1400]

- α equals the diffusion coefficient of X ,
- $b = 2p\alpha$ equals the constant drift of X , and
- $B(x) = \beta x + x\beta^\top$ equals the linear drift. Note that the drift enters (2.5) as its transpose $B^\top(u) = \beta^\top u + u\beta$.

By [2, Theorem 2.4] we have:

Proposition 2.4. *Let $\alpha \in S_d^+$ and $p \geq 0$.*

- If $p \geq \frac{d-1}{2}$, then for each $\beta \in M_d$, there exists a Wishart process on S_d^+ with parameters (p, α, β) .*
- Conversely, let X be a Wishart process with parameters (p, α, β) . Then $p \geq \frac{d-1}{2}$.*

We further recall the established fact [2, Theorem 2.7, equation (2.22)] that for each $x \in S_d^+$, (X, \mathbb{P}_x) can be realized as (a weak) solution of the stochastic differential equation

$$dX_t = \sqrt{X_t} dB_t Q + Q^\top dB_t^\top \sqrt{X_t} + (2p Q^\top Q + \beta X_t + X_t \beta^\top) dt \quad (2.6)$$

subject to $X_0 = x \in S_d^+$, for any $Q \in M_d$ which satisfies $Q^\top Q = \alpha$. Here B is a $d \times d$ standard Brownian motion, and \sqrt{X} denotes the unique matrix square root on the space of positive semi-definite matrices.

In the following we denote by ω_t^β the flow of the vector field $\beta x + x\beta^\top$, that is,

$$\omega^\beta : \mathbb{R} \times S_d^+ \rightarrow S_d^+, \quad \omega_t^\beta(x) := e^{\beta t} x e^{\beta^\top t}.$$

Its integral $\sigma_t^\beta : S_d^+ \rightarrow S_d^+$ for $t \geq 0$ is denoted by

$$\sigma^\beta : \mathbb{R}_+ \times S_d^+ \rightarrow S_d^+, \quad \sigma_t^\beta(x) = 2 \int_0^t \omega_s^\beta(x) ds.$$

Proposition 2.5. *Let $(X, \mathbb{P}_x)_{x \in S_d^+}$ be a Wishart process with parameter (p, α, β) . Then the characteristic exponents ϕ, ψ take the form*

$$\phi(t, u) = p \log \det \left(I + u \sigma_t^\beta(\alpha) \right), \quad (2.7)$$

$$\psi(t, u) = e^{\beta^\top t} \left(u^{-1} + \sigma_t^\beta(\alpha) \right)^{-1} e^{\beta t}. \quad (2.8)$$

Consequently, the Fourier-Laplace transform of X is given by

$$\mathbb{E}_x[e^{-\langle z, X_t \rangle}] = \left(\det(I + \sigma_t^\beta(\alpha) z) \right)^{-p} e^{-\text{tr} \left(z \left(I + \sigma_t^\beta(\alpha) z \right)^{-1} \omega_t^\beta(x) \right)}, \quad (2.9)$$

for all $z \in S_d^+ + iS_d$.

Proof. We first solve the generalized Riccati equations (2.4)–(2.5) for initial data $u \in S_d^+$. Formula (2.8) for ψ follows from the fact that $\frac{d}{dt} a^{-1}(t) = -a^{-1}(t) \frac{d}{dt} a(t) a^{-1}(t)$, see [4, Proposition III.4.2 (ii)]. Formula (2.7) follows by some elementary algebraic manipulations using the rule $\frac{d}{dt} \log(\det(a(t))) = \text{tr}(a^{-1}(t) \frac{d}{dt} a(t))$, see [4, Proposition II.3.3 (i)].

Concerning the Fourier-Laplace transform (2.9), we infer directly from their closed-form solutions (2.7)–(2.8) that $\phi(t, u)$ and $\psi(t, u)$ allow for analytic extensions to the complex tube $S_d^+ + iS_d$, which we denote by $\phi(t, z)$ and $\psi(t, z)$. Hence using analytic continuation, it can be seen that the Fourier-Laplace transform of X is given by

$$\mathbb{E}_x[e^{-\langle z, X_t \rangle}] = \left(\det(I + z \sigma_t^\beta(\alpha)) \right)^{-p} e^{-\text{tr} \left(\left(I + z \sigma_t^\beta(\alpha) \right)^{-1} z \omega_t^\beta(x) \right)}, \quad (2.10)$$

In order to obtain (2.9) from equation (3.5) it suffices to observe that for all $u, \theta \in M_d$ the following identity

$$u(I + \theta u)^{-1} = (I + u\theta)^{-1}u \quad (2.11)$$

holds, whenever either of both sides is well defined. \square

Combining Lemma 2.2 and Proposition 2.5, we obtain the following result concerning Wishart transition kernels and -densities:

Proposition 2.6. *Let $(X, \mathbb{P}_x)_{x \in S_d^+}$ be a Wishart process with parameters (p, α, β) . Suppose further that the diffusion parameter $\alpha \neq 0$. Then for each $(t, x) \in \mathbb{R}_{++} \times S_d^+$ $X_t^x \sim \Gamma(p, \omega_t^\beta(x); \sigma_t^\beta(\alpha))$. If $\alpha \in S_d^{++}$, $p > \frac{d-1}{2}$ and $t > 0$, then for all $x \in S_d^+$, X_t^x has a Lebesgue density*

$$f_t(x, \xi) = F(p, \omega_t^\beta(x), \sigma_t^\beta(\alpha), \xi).$$

2.3. On the central Wishart case. In the following we restate the characterization of the central Wishart laws by using [9]:

Theorem 2.7. *Let $d \geq 2$, $\sigma \in S_d^{++}$ and $p \geq 0$. The following are equivalent:*

- (i) *Formula (1.2) is the Laplace transform of a probability measure $\Gamma(p, \omega; \sigma)$ on S_d^+ .*
- (ii) *$p \in \Lambda_d$.*

Proof. (i) \Rightarrow (ii): This is [9, Theorem 1], as we exclude the point mass at 0. The converse direction is a special case of Lemma 2.2 (ii) and (iii). \square

It is important to note that condition (ii) is not necessary, if σ is degenerate. In fact, it is easy to prove by use of orthogonal transformations (see <http://arxiv.org/abs/1009.3708>)

Corollary 2.8. *Let $r = \text{rank}(\sigma)$. The following are equivalent:*

- (i) *Formula (1.2) is the Laplace transform of a probability measure $\Gamma(p, \omega; \sigma)$ on S_d^+ .*
- (ii) *$p \in \Lambda_r$.*

As a trivial consequence, one has

Corollary 2.9. *The following are equivalent:*

- (i) *$\Gamma(p; \sigma)$ is infinitely divisible.*
- (ii) *$\text{rank}(\sigma) = 1$.*

3. PROOF OF THEOREM 1.1

Let $(p, \omega, \sigma) \in \mathbb{R}_{++} \times S_d^+ \times S_d^{++}$ such that $\mu := \Gamma(p, \omega; \sigma)$ is a probability measure, that is, eq. (1.1) holds. The domain of its moment generating function is defined as

$$D(\mu) := \{u \in S_d \mid \mathcal{L}_\mu(u) := \int_{S_d^+} e^{-\langle u, \xi \rangle} \mu(d\xi) < \infty\},$$

which is the maximal domain to which the Laplace transform, originally defined for $u \in S_d^+$ only, can be extended. It is well known that $D(\mu)$ is a convex (hence connected) set, and we also know that $S_d^+ \subset D(\mu)$. Clearly $(I + \sigma u)$ is invertible if and only if the (symmetric) matrix $(I + \sqrt{\sigma}u\sqrt{\sigma})$ is non-degenerate. Using these facts and the defining equation (1.1) we infer that

$$D(\mu) := \{u \in S_d \mid (I + \sqrt{\sigma}u\sqrt{\sigma}) \in S_d^{++}\} = -\sigma^{-1} + S_d^{++}, \quad (3.1)$$

and therefore $D(\mu)$ is even open. Accordingly, the natural exponential family of μ is the family of probability measures⁴

$$F(\mu) = \left\{ \frac{\exp(v\xi)\mu(d\xi)}{\mathcal{L}_\mu(v)} \middle| v \in -\sigma^{-1} + S_d^{++} \right\}.$$

We start by stating some key properties of Wishart distributions⁵:

Proposition 3.1. (i) Let $p \geq 0$, $\omega \in S_d^+$. Suppose X is an S_d^+ -valued random variable distributed according to $\Gamma(p, \omega; I)$. Let $q \in S_d^+$ and set $\sigma := q^2$. Then $qXq \sim \Gamma(p, q\omega q; \sigma)$ ⁶. In particular, $\Gamma(p, \omega; I)$ exists if and only if $\Gamma(p, q\omega q; \sigma)$ exists.
(ii) Let $p \geq 0$, $\sigma \in S_d^{++}$ and $\omega \in S_d^+$ such that $\mu := \Gamma(p, \omega; I)$ is a probability measure. For $v = \sigma^{-1} - I$ we have that

$$\frac{\exp(v\xi)\mu(d\xi)}{\mathcal{L}_\mu(v)} \sim \Gamma(p, \sigma\omega\sigma; \sigma). \quad (3.2)$$

Conversely, if $\Gamma(p, \sigma\omega\sigma; \sigma)$ is a well defined probability measure, so is μ , and (3.2) holds. In particular, we have that the exponential family generated by μ is a Wishart family and equals

$$F(\mu) = \{\Gamma(p, \sigma\omega\sigma, \sigma) \mid \sigma \in S_d^{++}, \quad \sigma^{-1} - I \in D(\mu)\}.$$

(iii) Let $\Gamma(p, \omega_0; \sigma_0)$ be a probability measure, where $\sigma_0 \in S_d^{++}$. Then we have
(a) $\Gamma(p, t\omega_0; \sigma_0)$ is a probability measure for each $t > 0$.
(b) If, in addition, ω_0 is invertible, then $\Gamma(p, \omega; \sigma)$ is a probability measure for each $\omega \in S_d^+$, $\sigma \in S_d^+$.

Proof. Let \mathbb{E} be the corresponding expectation operator. By repeated use of the cyclic property of the trace and by the product formula for the determinant, we have

$$\begin{aligned} \mathbb{E}[e^{-\langle u, qXq \rangle}] &= \mathbb{E}[e^{-\langle quq, X \rangle}] = \det(I + quq)^{-1} \exp(-\text{tr}(quq(I + quq)^{-1}\omega)) \\ &= \det(I + \sigma u)^{-1} \exp(-\text{tr}(uq(I + quq)^{-1}q^{-1}q\omega q)) \\ &= \det(I + \sigma u)^{-1} \exp(-\text{tr}(u(I + \sigma u)^{-1}q\omega q)), \end{aligned}$$

which proves assertion (i). Next we show (ii). We note first, that by (3.1) we have that $v = \sigma^{-1} - I \in D(\mu)$. Hence exponential tilting is admissible. Furthermore, we have

$$\int_{S_d^+} e^{-\langle u+v, \xi \rangle} \Gamma(p, \omega; I)(d\xi) = \det(1 + (u + v))^{-p} \exp(-\text{tr}((u + v)(1 + u + v)^{-1}\omega)), \quad (3.3)$$

and setting $v = \sigma^{-1} - I$ we obtain

$$1 + u + v = \sigma^{-1}(1 + \sigma u).$$

Hence the first factor on the right side of eq. (3.3) is proportional to $\det(1 + \sigma u)^{-p}$. It remains to show that

$$-\text{tr}((u + v)(1 + u + v)^{-1}\omega) = c + \text{tr}(u(1 + \sigma u)^{-1}\sigma\omega\sigma) \quad (3.4)$$

⁴In order to avoid confusions with calculations in the proof of the upcoming proposition, we change here from u notation to v , because u denotes the Fourier-Laplace variable in this paper.

⁵Some related properties can be found in Letac and Massam [7], but in a different notation. More detailed information may be found in Appendix A

⁶Expressed in geometric language, we say that the pushforward of $\Gamma(p, \omega; I)$ under the map $\xi \mapsto q\xi q$ equals $\Gamma(p, q\omega q; \sigma)$.

for some real constant c , because then the right side of (3.3) is proportional to the Laplace transform of $\Gamma(p, \sigma\omega\sigma; \sigma)$. To this end, we do some elementary algebraic manipulations:

$$\begin{aligned}
-(u+v)(I+u+v)^{-1}\omega &= -(u-1+\sigma^{-1})[\sigma^{-1}(1+\sigma u)]^{-1}\omega \\
&= -(-1+\sigma^{-1}+u)(\sigma^{-1}+u)^{-1}\omega \\
&= -\omega + (\sigma - \sigma)\omega + (\sigma^{-1}+u)^{-1}\omega \\
&= (\sigma - I)\omega - \sigma(\sigma^{-1}+u)(\sigma^{-1}+u)^{-1}\omega + (\sigma^{-1}+u)^{-1}\omega \\
&= (\sigma - I)\omega - \sigma u(\sigma^{-1}+u)^{-1}\omega \\
&= (\sigma - I)\omega - \sigma u(I + \sigma u)^{-1}\sigma\omega.
\end{aligned}$$

We set now $c := \text{tr}((\sigma - I)\omega)$ which is the real number we talked about before. Taking trace and performing cyclic permutation inside, we obtain (3.4), and therefore the identity (3.2) is shown. The assertion concerning the exponential family follows by the very definition of the latter.

We may therefore proceed to (iii) which is proved by repeatedly applying (i) and (ii): Let $\Gamma(p, \omega_0; \sigma_0)$ be a probability measure. Then by (ii), also $\Gamma(p, \sigma_0^{-1}\omega_0\sigma_0^{-1}; I)$ is one. Let q_1 such that $q_1^2 = \sigma_1 \in S_d^{++}$. We may write $\Gamma(p, \sigma_0^{-1}\omega_0\sigma_0^{-1}; I) = \Gamma(p, q_1^{-1}(q_1\sigma_0^{-1}\omega_0\sigma_0^{-1}q_1)q_1^{-1}; I)$, and by applying (i), we obtain the pushforward measure $\Gamma(p, q_1\sigma_0^{-1}\omega_0\sigma_0^{-1}q_1; \sigma_1)$. By (ii) we have that $\Gamma(p, q_1^{-1}\sigma_0^{-1}\omega_0\sigma_0^{-1}q_1^{-1}; I)$ is a probability measure as well, and once again by (ii) we infer that for all $\sigma \in S_d^{++}$, $\Gamma(p, \sigma q_1^{-1}\sigma_0^{-1}\omega_0\sigma_0^{-1}q_1^{-1}\sigma; \sigma)$ is a probability. We use this fact to prove both parts of the assertion. Without loss of generality we assume that σ is non-degenerate, because in the case $\sigma \in \partial S_d^+$ we may invoke Lévy's continuity theorem⁷. Setting $q_1 = 1/\sqrt{t}I$ and $\sigma = \sigma_0$, we see that (iii)a holds. For $\omega_0 \in S_d^+$ we choose $q_1 \in S_d^+$ such that $q_1^{-1}\sigma_0^{-1}\omega_0\sigma_0^{-1}q_1^{-1} = \sigma^{-1}\omega\sigma^{-1}$, which allows to conclude (iii)b. \square

Finally, we deliver our proof of Theorem 1.1:

Proof. Let $p > 0$ such that for some $\omega_0 \in S_d^+$, $\sigma \in S_d^{++}$, the right side of (1.1) is the Laplace transform of a non-trivial probability measure $\Gamma(p, \omega_0; \sigma)$. By Proposition 3.1 (iii)a, we have that $\Gamma(p, \omega_0/n; \sigma)$ is a probability measure for each $n \in \mathbb{N}$. Letting $n \rightarrow \infty$ and invoking Lévy's continuity theorem, we obtain that $\Gamma(p; \sigma)$ is a probability measure. But then by the characterization of central Wishart laws, Theorem 2.7 (ii), we have that $p \in \Lambda_d$.

Let now $p_0 \in \Lambda_d \setminus [\frac{d-1}{2}, \infty)$, and let us assume, by contradiction, that there exist $(\omega_0, \sigma) \in S_d^+ \times S_d^{++}$, $\text{rank}(\omega_0) > 2p_0 + 1$ such that $\Gamma(p_0, \omega_0; \sigma)$ is a probability measure. Pick now $\omega_1 \in S_d^+$ such that $\omega^* := \omega_1 + \omega_0$ has $\text{rank}(\omega^*) := \text{rank}(\omega_1) + \text{rank}(\omega_0) = d$, and set $p_1 := \frac{d - \text{rank}(\omega_0)}{2}$. By construction $2p_1 = \text{rank}(\omega_1)$, and $p_1 \in \Lambda_d \setminus [\frac{d-1}{2}, \infty)$. Hence Proposition 2.2 (ii) implies the existence of a non-central Wishart distribution $\Gamma(p_1, \omega_1, \sigma)$. Note that $p^* := p_0 + p_1 \in \Lambda_d \setminus [\frac{d-1}{2}, \infty)$ and that by convolution

$$\Gamma(p^*, \omega^*, \sigma) := \Gamma(p_0, \omega_0, \sigma) \star \Gamma(p_1, \omega_1, \sigma)$$

⁷Strictly speaking, Lévy's continuity theorem applies to characteristic functions. However, in the Wishart case, the right side of (1.1) can even be extended to even the Fourier-Laplace transform with ease, and by preserving its functional form.

is a probability measure as well. Since ω^* is of full rank, we have by Proposition 3.1 (iii)b that $\Gamma(p^*, \omega; \sigma)$ is a probability measure for all $(\omega, \sigma) \in (S_d^+)^2$. Hence $\Gamma(p^*, \omega; t\sigma)$ exists for all $(t, \omega, \sigma) \in \mathbb{R}_+ \times (S_d^+)^2$.

We proceed by reverse engineering of the results of section 2.2. Pick any $\alpha \in S_d^+ \setminus \{0\}$. For each $(t, x) \in \mathbb{R}_+ \times S_d^+$ we let $p_t(x, d\xi)$ be the probability measure given by the Laplace transform

$$\int_{S_d^+} e^{-\langle u, \xi \rangle} p_t(x, d\xi) = (\det(I + 2t\alpha u))^{-p^*} e^{\text{tr}(-u(I+2t\alpha u)^{-1}x)}, \quad (3.5)$$

(cf. (2.9) for $\beta = 0$). By the proof of Proposition 2.5 we know that $\phi(t, u) := p^* \log(I + 2t\alpha u)$ and $\psi(t, u) := u(I + 2t\alpha u)^{-1}$ satisfy the system of Riccati equations (2.4)–(2.5) with $\beta = 0$. From a density argument it follows that the function $p_t(x, d\xi)$ satisfies the Chapman-Kolmogorov equation, hence it is the transition function of a Markov process X on S_d^+ and by construction the Laplace transform is exponentially affine in the state variable x . Hence X is an affine process in the sense of [2], with constant drift parameter $b = 2p^*\alpha$ and diffusion coefficient α . But $b = 2p^*\alpha \not\geq (d-1)\alpha$, which contradicts the drift condition formulated in Proposition 2.4 (ii). Therefore $\text{rank}(\omega_0) \leq 2p_0 + 1$ and we have proved the first part of the theorem.

The second part of the theorem follows from Lemma 2.2 (ii) and (iii). \square

APPENDIX A. REMARKS ON ALTERNATIVE DEFINITIONS OF THE WISHART DISTRIBUTIONS

A number of different notations and definitions of Wishart distributions appear in the literature. This paper uses several technical tools which are presented in Letac and Massam's work [7], and therefore we have chosen a notation which is closely related to the latter.

Letac and Massam use instead of $\Gamma(p, \omega; \sigma)$ the parameterized family $\gamma(p, a; \sigma)$, where ω is replaced by $a := \sigma^{-1}\omega\sigma^{-1}$. Accordingly (1.1) can be written in the form

$$\mathcal{L}(\gamma(p, a; \sigma))(u) = (\det(I + \sigma u))^{-p} e^{-\text{tr}(u(I + \sigma u)^{-1}\sigma a \sigma)}, \quad u \in S_d^+. \quad (\text{A.1})$$

Note that this requires σ to be invertible. Other authors use densities to define Wishart distributions. There are, however, two notable disadvantages of using densities rather than the Laplace transform or the characteristic function:

- A density need not always exist: If σ is degenerate, $\Gamma(p, \omega; \sigma)$ is not absolutely continuous with respect to the Lebesgue measure on S_d^+ . To see this we assume for a contradiction that $\Gamma(p, \omega; \sigma)$ has a Lebesgue density, for some σ of rank $r < d$. Let X be an S_d^+ -valued random variable distributed according to $\Gamma(p, \omega; \sigma)$. Since linear transformations do not affect the property of having a density and since the non-central Wishart family is invariant under linear transformations (this is easy to check), we may without loss of generality assume that $\sigma = \text{diag}(0, I_r)$, where I_r is the $r \times r$ unit matrix. Consider the projection

$$\pi_r : x = (x_{ij})_{1 \leq i, j \leq d} \mapsto \pi_r(x) := (x_{ij})_{1 \leq i, j \leq r}.$$

A simple algebraic manipulation yields that the Laplace transform of $\pi_r(X)$ equals

$$e^{-\text{tr}(\pi_r(\omega)v)}, \quad v \in S_r^+,$$

which is the Laplace transform of the unit mass concentrated at $\pi_r(\omega)$. But the pushforward of a measure with density under a projection must have a density again. This yields the desired contradiction.

- The density for non-central Wishart distributions is complicated, as it is a series expansion in zonal polynomials, see [7].

As we do not work with the explicit form of the densities of non-central Wishart distributions, we did not need to specify the precise form of the latter in Lemma 2.2 (iv).

Our use of the Laplace transform throughout the paper is due to the use of affine processes, which is a class of Markov processes with the key defining property that their Fourier-Laplace transform is of exponentially affine form in the state variable (see, subsection 2.2, in particular the defining equation (2.1)). One of the important consequences of this property is that the affine exponents can be determined by solving a system of ordinary differential equations, the so-called Riccati Differential Equations (2.4)–(2.5), rather than the Forward–Kolmogorov–PDE for the Laplace transform,

$$\partial_t \Phi(t, u, x) = \mathcal{A}\Phi(t, u, x), \quad \Phi(0, u, x) = \exp(-\text{tr}(ux)),$$

where \mathcal{A} denotes the generator of the affine process. In other words, the last equation can be solved by an Ansatz of the form $\Phi(t, u, x) := \exp(-\phi(t, u) - \text{tr}(\psi(t, u)x))$, as it leads to the Riccati ODEs and let us avoid solving for complicated parabolic PDEs. For more technical details and a complete theory of matrix-variate affine processes we refer the interested reader to [2].

Even though specifying Wishart distributions by means of their densities is very natural, the best known construction is by pushforwards of normal distributions under certain quadratic forms: Let $\xi_1, \xi_2, \dots, \xi_k$ be a sequence of \mathbb{R}^d -valued normally distributed random variables with mean vectors $\mu_i \in \mathbb{R}^d$ and covariance matrix Σ . By using Lemma 2.2 (i) we infer that the S_d^+ -valued random variable $\Xi := \sum_{i=1}^k \xi_i \xi_i^\top$ has distribution $\Gamma(p, \omega; \sigma)$, where $p = 2k$, $\omega = \sum_{i=1}^k \mu_i \mu_i^\top$ and $\sigma = 2\Sigma$.

In Gupta and Nagar's notation [6] this reads as follows. We define the $d \times k$ matrix $M = (\mu_1, \dots, \mu_k)$ and the $d \times d$ matrix $\Theta := \Sigma^{-1} M M^\top$. Using the matrix-variate normal distribution, we have a $d \times k$ matrix-valued random variable $X := (\xi_1, \xi_2, \dots, \xi_k)$ which is distributed according to $\mathcal{N}_{d,k}(M, \Sigma \otimes I)$, and

$$X X^\top \equiv \Xi \sim \mathcal{W}_d(k, \Sigma, \Theta), \quad (\text{A.2})$$

where k is the shape parameter, Σ is the scale parameter, and the parameter of non-centrality equals Θ ([6, Theorem 3.5.1]). As with Letac and Massam's class of generalized non-central distributions, this imposes that Σ must be invertible. Accordingly, the Laplace transform of Ξ is given by [6, Theorem 3.5.3]

$$\mathbb{E}[e^{-\text{tr}(u\Xi)}] = \det(I + 2\Sigma u)^{-k/2} e^{-\text{tr}(\Theta(I + 2\Sigma u)^{-1} \Sigma u)}. \quad (\text{A.3})$$

When $\Sigma = I$ then we see that $\mathcal{W}_d(k, \Sigma, \Theta = M M^\top)$ equals $\Gamma(k, M M^\top; 2\Sigma)$. Let us now have a look at the general case, where $\Sigma \neq I$ and where Θ can be any positive semidefinite matrix. We observe that the right side of eq. (A.3) can be also written in the form

$$\mathbb{E}[e^{-\text{tr}(u\Xi)}] = \det(I + 2u\Sigma)^{-k/2} e^{-\text{tr}(\Theta u \Sigma (I + 2u\Sigma)^{-1})},$$

which is the notation found in [9, equation (2)]. To see this, one may use for the first factor the multiplicativity of the determinant, while the second factor follows from the identity

$$[(I + 2\Sigma u)^{-1} \Sigma u]^{-1} = (\Sigma u)^{-1} (I + 2\Sigma u) = u^{-1} \Sigma^{-1} + 2I$$

as well as

$$[u \Sigma (I + 2u\Sigma)^{-1}]^{-1} = \Sigma^{-1} u^{-1} + 2I$$

and the symmetry of Θ, Σ and u .

We finally show that $\mathcal{W}_d(k, \Sigma, \Theta) = \Gamma(k/2, (\Theta\sigma + \sigma\Theta)/4; \sigma)$. The exponent on the right hand side of eq. (A.3) can be rewritten as

$$\begin{aligned} \text{tr}(\Theta(I + \sigma u)^{-1}\sigma u/2) &= \text{tr}((\sigma u)^{-1} + \sigma^{-1})^{-1}\Theta) = \text{tr}((u^{-1} + \sigma^{-1})^{-1}\Theta\sigma/2) \\ &= \text{tr}((u^{-1} + \sigma^{-1})^{-1}\frac{\Theta\sigma + \sigma\Theta}{4}) \end{aligned}$$

and since Θ, Σ are positive semidefinite, also $\Theta\sigma + \sigma\Theta$ is positive semidefinite. However, for invertible Σ , the map $X \mapsto \Sigma X + X\Sigma$ is injective⁸, but not surjective⁹, in general (unless Σ is a multiple of the unit matrix). Hence the class of non-central Wishart distributions used in this paper is strictly larger than the classes in the standard literature, if we insist on positive semidefinite non-centrality parameters Θ . If we, however, allow (not necessarily positive semidefinite) non-centrality parameters of the form $\Theta = 2\sigma^{-1}\omega$, where $\omega \in S_d^+$, then all the three mentioned Wishart classes coincide. This definition is also in line with the quadratic construction of Ξ , see eq. (A.2). Note, however, that [9] imposes a symmetric non-centrality parameter, which means that their class of Wishart distributions is strictly smaller than ours. Hence Corollary 1.2 comprises a more general situation than [9, Theorem 2].

REFERENCES

- [1] M.-F. BRU, *Wishart Processes*, J. Theoret. Probab., 4 (1991), pp. 725–751.
- [2] C. CUCHIERO, D. FILIPOVIĆ, E. MAYERHOFER, AND J. TEICHMANN, *Affine processes on positive semidefinite matrices*. Ann. Appl. Probab., 21 (2), (2011), pp. 397–463.
- [3] D. DUFFIE, D. FILIPOVIĆ, AND W. SCHACHERMAYER, *Affine Processes and Applications in Finance*, Ann. Appl. Probab., 13 (2003), pp. 984–1053.
- [4] J. FARAUT AND A. KORÁNYI, *Analysis on symmetric cones*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1994. Oxford Science Publications.
- [5] S. G. GINDIKIN, *Invariant generalized functions in homogeneous domains*, J. Funct. Anal. Appl. 9 (1975) 50–52.
- [6] A. K. GUPTA AND D. K. NAGAR, *Matrix variate distributions*, vol. 104 of Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, Chapman & Hall/CRC, Boca Raton, FL, 2000.
- [7] G. LETAC AND H. MASSAM, *The noncentral Wishart as an exponential family, and its moments*, J. Multivariate Anal., 99 (2008), pp. 1393–1417.
- [8] P. LÉVY, *The arithmetical character of the Wishart distribution*, Proc. Cambridge Philos. Soc., 44 (1937).
- [9] S. D. PEDDADA AND D. ST. P. RICHARDS, *Proof of a conjecture of M. L. Eaton on the characteristic function of the Wishart distribution*, Ann. Probab., 19 (1991), pp. 868–874.
- [10] D. SHANBHAG, *The Davidson-Kendall Problem and related results on the structure of the Wishart distribution*, Austr. J. Statist., 30A (1988), pp. 272–280.
- [11] J. WISHART, *The generalised product moment distribution in samples from a normal multivariate population*, Biometrika, 20A (1928), pp. 32–52.

DEUTSCHE BUNDESBANK, RESEARCH CENTRE, WILHELM-EPSTEIN-STR.14, 60431 FRANKFURT AM MAIN, GERMANY

E-mail address: eberhard.mayerhofer@gmail.com

⁸This can be seen by diagonalizing Σ .

⁹It is well known that any linear transformation on S_d^+ is of the form $X \mapsto gXg^\top$, where g is a real, not necessarily symmetric, $d \times d$ matrix.